Tracking of Tissue Movement Using Distance-Weighted Log Ratio Similarity Matching Algorithm

Rohana Abdul Karim[1,2], Mohd Marzuki Mustafa[2], Mohd Asyraf Zulkifley[2]
[1]*Faculty of Electrical & Electronics Engineering, Universiti Malaysia Pahang Kampus Pekan, 26600 Pekan, Pahang, Malaysia*
[2]*Department of Electric, Electronic & Systems Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia*
rohanaak@ump.edu.com

***Abstract*—Nowadays, the growth of health care quality awareness lead to the advancement of the medical technologies, especially for surgery technologies. In the field of computer vision, tracking of the tissues and internal organs (TDOD) movements have been beneficial to many surgical technologies such as computer-assisted surgery and minimally invasive surgery. TDOD tracking poses a challenging task due to the nature characteristic of TDOD which mainly has a homogenous surface and texture. We proposed a feature point tracking algorithm based on hypothesis testing t-test as a novel technique for TDOD tracking. This algorithm is based on the distance-weighted log ratio t-test similarity measurement. The algorithm has been tested and showed it can perform better compared with existing methods in all the test datasets.**

***Index Terms*—Tissues and Internal Organs; Feature Point; Matching; Hypothesis Test; Distance.**

I. INTRODUCTION

Image registration technique is a core step in tracking objects that have been moved from its original location. The mapping process is essential for providing coordinates information for objects of interest. Such information is used for high-level image processing such as reconstruction image from 2-dimensional to 3-dimensional images [1] and tracking the trajectory changes [2][3].

There are two types of tracking methods that are often used in image processing: 1) mathematical modeling and 2) feature point matching. For mathematical modeling, the matrix transformation technique is one of the popular methods for mapping 2-dimensional images [4][5]. Initially, a few corresponding objects are identified. Information from the corresponding objects is then used to determine the unknown matrix transformation parameters such as angle, translation and scale. By using all the extracted information, new coordinates for each pixel in the whole image can be estimated. Many studies have shown that the transformation matrix has achieved a good performance for rigid objects with uniform movements [6][7].

However, the nature of TDOD which is non-rigid and moves dynamically leads to the complex processing [8]. There are two factors affecting the TDOD movement; 1) natural movement, and 2) interruption from external. Natural movement refers to the ability of TDOD to move and change the size and form unconsciously such as an expansion and contraction of the cardiac and lung during the breathing process. Furthermore, the soft texture of the TDOD surfaces causes the shape and direction of movement to be easily affected whenever there is interference from external factors such as due to surgical instruments. Consequently, it is a challenging task to track the TDOD movement.

Recently, image registration based on feature point matching has been applied for tracking TDOD movements [9][10][11][12]. The feature point matching works by searching the similar feature point between two scenes of images based on the feature descriptor information. The feature point is a good and unique feature in the image. A good feature point should enable the point to be repeatedly detected even though there are changes in the view, translation, rotation, scaling and pixels intensities [13][14].

In other fields, the statistical hypothesis test has been widely used in the field of geochemical, geology and biomedical for searching the similarities between two substances. For example, in the field of geochemical, the hypothesis test is used to find out the similarity in a mineral composition in the soil. Reimann and Caritat figure [15] have applied the hypothesis test to build up the geochemical map. The map has information regarding a mineral similarity between Europe and Australia continent for a better understanding of the demography of the natural mineral composition. The same concept has also been used to find out the similarity of chemical contents between the sample materials with the standard chemical materials [16][17].

In geology, the hypothesis test has been used to determine the existence of minerals in the soil [18]. The sediment samples are analyzed by comparing the composition of the sample data with the standard reference sediment that does not contain any mineral. The difference between the sample sediment and the standard reference indicates the presence of minerals. Similarly, the same technique for mineral search is also utilized in agriculture [19][20] and geostatistical [21] to determine the mineral contents in the soil for agricultural areas.

Another example of the use of hypothesis testing in the biomedical field is to study the cause and effect of a disease such as epidemiology [22]. Epidemiology is the study to understand how and why the diseases occurred in different groups of people. This epidemiological information is then used to plan and evaluate strategies to prevent and draw guidelines for managing infected patients.

Based on the success of hypothesis testing for searching

similarities in the other fields, we are motivated to investigate the similar approach on matching feature points in TDOD. Our works on the hypothesis testing applied on the tracking of TDOD feature points has been reported in [23]. In this paper, we will describe the improvement on this method by incorporating distant criteria.

II. EXPERIMENTAL

The tracking algorithm consists of three main components: 1) feature extraction, 2) feature description, and 3) feature matching. Feature extraction is to extract the keypoints in the images by using centers surrounded by extreme (STAR) detector [24]. Next, feature descriptor module generates the unique signatures' for each keypoint by using vector set of log ratio. The unique signature enables the keypoint to be differentiated with other keypoints. The descriptor uses image intensities solely as a feature and transforms it into log ratio. The last process of the algorithm is feature matching which is based on t-test hypothesis testing. Both feature descriptor and feature matching algorithms have been discussed in our previous worked [23]. An overview of our algorithms is depicted in Figure 1. The details of each process are explained as below:

1. LRD: is a subtraction between interest keypoint feature descriptor and candidates' keypoint feature descriptor. Candidates' keypoints are obtained from ten nearest neighbours based on the last known location of interest keypoint.

$$LRD = V_i - V_c \quad (1)$$

where V_i is a set vector of interest keypoint and V_c is a set vector of candidates' keypoint.

2. Normality test: is a crucial step for t-test analysis because t-test assumes all the data follow the normal distribution. Therefore, this step filters out abnormal keypoints. There are two types of normality test that are appropriate to be used which are; Lilliefors and Jarque-Bera tests. We tested for the normality based on region as illustrated in Figure 2. In normal behavior, TDOD moves in a small displacement [25], [26]. Therefore, the nearest keypoints from the previous location of keypoint of interest location (region A) have a higher likelihood to be the correct matched keypoint. Because of this higher likelihood, a lower threshold for the significant value is used to test the candidate keypoints in this region. Meanwhile, keypoints in region B which are further from previously matched keypoint must be screened more tightly to ensure only very credible candidates' keypoints are listed as normal.

3. Paired t-test: This step measures the similarity probability between the keypoint of interest with every keypoint candidates. Based on 95% of significant value [27][28], we define P1 as the probability of true mean LRD, which is less than the lower confidence interval. Meanwhile, P2 is the probability of mean LRD that is greater than the upper confidence interval. The acceptable probability (PA) is calculated as;

$$PA = \big((P1 + P2)\big|PA \geq 0.05\big) \quad (2)$$

The keypoint with maximum acceptable probability is selected as final matched keypoint. To further improve the matching accuracy, we have incorporated two distance-weighted criteria.

A. Threshold Distance (THD)

The aim of threshold distance is to filter out improbable keypoint candidates'. Any keypoint candidates that exceed the threshold T_d is classified as improbable and then removed from the candidate lists. Finally, the candidate within the threshold with the highest acceptable probability is selected as a final matched keypoint.

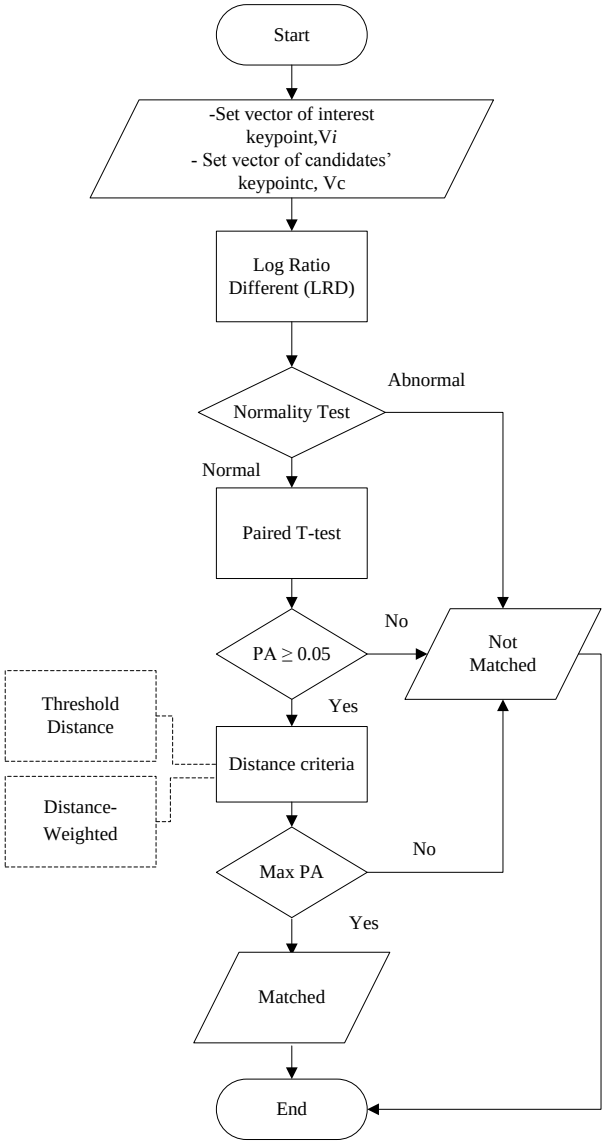


Figure 1: The overview of the matching algorithm by using hypothesis testing. Dash line indicates the proposed modules

B. Distance-Weighted (DW)

Instead of using single region, the second approached dividing the distance into multiple regions. We classify the tendency of the match into three categories; 1) very highly matched region (A_w), 2) declining matching region (B_w), and 3) unlikely matched region (C_w) as shown in Figure 3. The DW weightage is calculated by multiply the weighted distance factor (y) with acceptable probability.

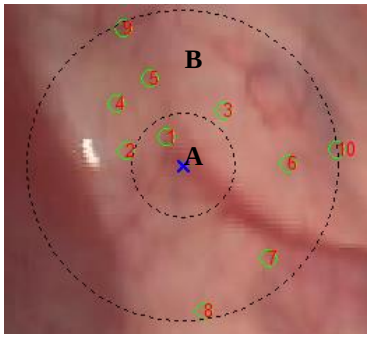


Figure 2: Regions for normality test. Blue 'x' is the previous location of keypoint of interest.

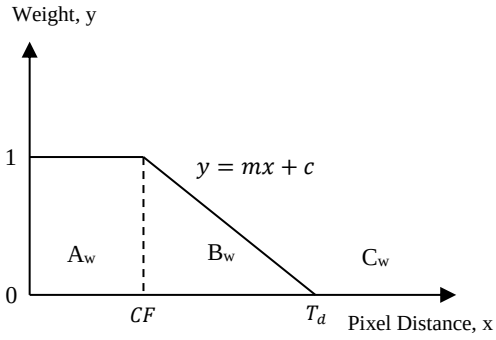


Figure 3: An illustration of distance-weighted criteria

The DW matching algorithm is started by calculating the distance (x) of candidate keypoints from the previous location of keypoint of interest. Based on the distance, candidate keypoints are weighted according to the region criterion. For region A, the distance range starts from 0 to cut off distance, CF. In this region, each candidate keypoint has the same weight. Due to that, equal weights (y) are given to the candidate keypoints which located within the region A. The new PA_w score is calculated by Equation (3) with y equal to 1.

$$PA_w = PA * y \tag{3}$$

For region B, the distance range is between $CF < x \leq T_d$. Each candidate keypoint in this region has a different weight, which is depending on their distance. As the distance increases, the weighting factor is slightly decreased. The decrement indicates that the candidate keypoints are slowly lost their tendencies to be matched. Within this range, the weighting factor is modelled by linear equation;

$$y = mx + c \tag{4}$$

The last region C is, where the distance range is greater than T_d pixels $x > T_d$. All candidate keypoints within this range are assumed as improbable to be matched since large movement is unlikely. Hence, at this region, the weighting factor is equal to zero. Overall, the DW model can be summarized as;

$$PA_w = PA * y \begin{cases} 0 \leq x \leq CF & y = 1 \\ CF < x \leq T_d & y = mx + c \\ x > T_d & y = 0 \end{cases} \tag{5}$$

III. Results And Discussions

A. Data Set and Ground Truth

Our algorithm was tested using liver, heartbeat and abdomen movements videos obtained from the public database [31]. Each data set consisted of 30 frames with a sampling interval of three frames for motion analysis. The dataset represented various types of movements. There was translation, rotation and scaling due to camera and natural organ tissues movements. The ground truth has been determined manually by the researchers.

B. Performance measurement

The effectiveness of our algorithms was evaluated by using mean error distance (MED), which is defined as;

$$MED = \frac{\sum_1^N M(x,y) - GT(x,y)}{N} \tag{6}$$

where $M(x,y)$ is final matched keypoints obtained from proposed algorithms and $GT(x,y)$ is the ground truth location. N denotes the total numbers of final matched keypoints for each frame.

C. Parameters Setup

There are two parameters derived from the analysis of TDOD matching algorithm without distance criteria as depicted in Figure 4, which are T_d and CF. The T_d parameter was calculated from the mean of 75% percentile of all data set which is equal to 15 pixels. However, there had some cases where the MED of intended match keypoint increased to 20 pixels. As a precaution step, the T_d value is set to 20 pixels. Meanwhile, the value of CF parameter is defined as;

$$CF = \frac{\sum_1^v BPi_{min}}{v} \tag{7}$$

where BPi is a 25% percentile of the box plot for each data set and v is a total number of data set. Thus, the CF value is equal to 4.5 pixels.

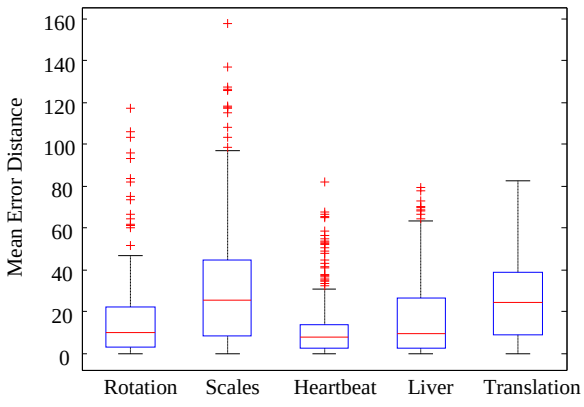


Figure 4: Box plot analysis for hypothesis testing as a TDOD matching algorithm without distance criteria.

D. Results

1) Normality Test: Single Test Vs Multiple Test

This experiment was performed with the aim of finding out the most appropriated test for normal distributions technique. Table-1 showed that the combination of Lilliefors test and

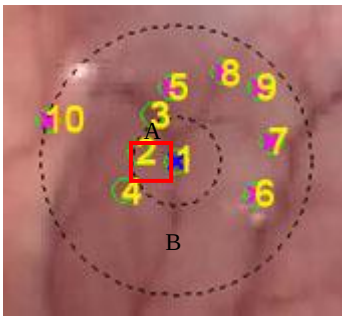
Jarque-Bera test managed to increase the accuracy of the candidates' keypoints by achieving the lowest value of MED.

Table 1
Comparison of MED Performance Between Single Normality Test and Combined Normality Test.

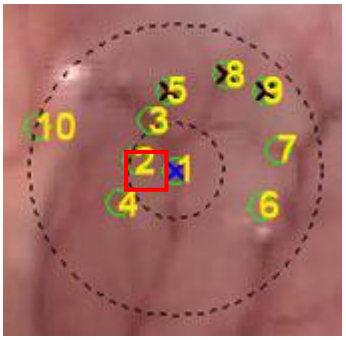
Movement Types	MED		
Data set	Lilliefors	Jarque-Bera	Lilliefors + Jarque-Bera
Rotation	18.32	18.10	14.83
Scales	28.47	35.37	22.13
Heartbeat	14.81	15.95	13.32
Liver	17.81	19.28	14.92
Translation	20.58	24.21	17.31
Average	20.00	22.58	16.50

2) Matching Performance

Figure 5 shows a comparison of final matched keypoint for one sample of keypoint of interest. The threshold distance model was useful for shortlisting the best candidate keypoints. However, in most cases, threshold distance was incapable of selecting the intended keypoint (point no.1 in Figure 5) as final matched keypoint because there have another candidate keypoint obtained the highest PA as shown in Figure 5(a). In contrast, the DW model did not prune any candidate keypoints. But, the DW model weighted the similarity probability of each candidate keypoints according to their distance respectively. Figure 5(b) shows that the intended keypoint obtained the highest probability by using DW algorithm. Meanwhile, Figure 6(a)-(b) display a visual performance for final matched keypoints.



(a) Threshold Distance

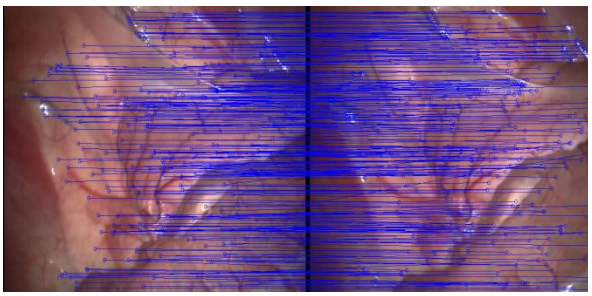


(b) Distance-Weighted

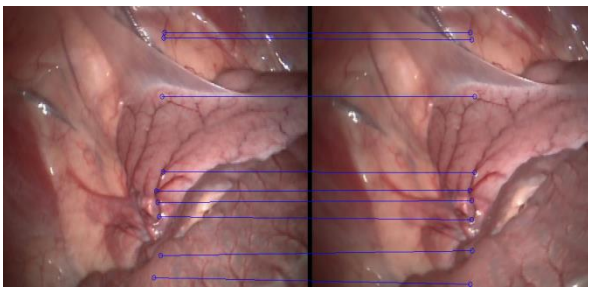
Figure 5: Sample of final matched keypoint for one sample of keypoint of interest.

Table 2
Samples of The Probability Acceptance for Candidate Keypoint Using THD and DW Techniques.

Keypoint No.	PA (THD)	Distance (Pixel)	PA (DW)
1	0.296	0	0.296
2	0.295	11.40	0.169
3	0.432	15.65	0.138
4	0.364	18.35	0.057
6	-	26	0
7	-	29.61	0
10	-	40.80	0



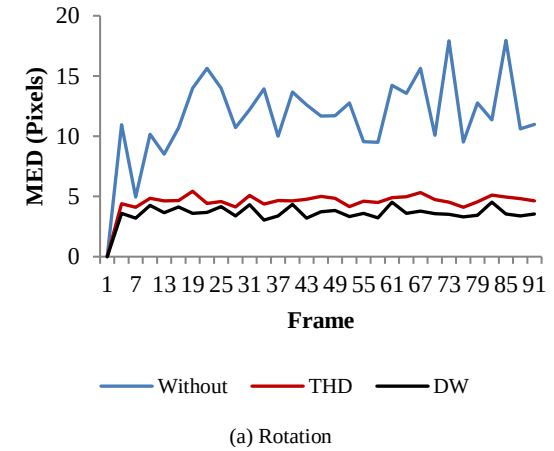
(a)



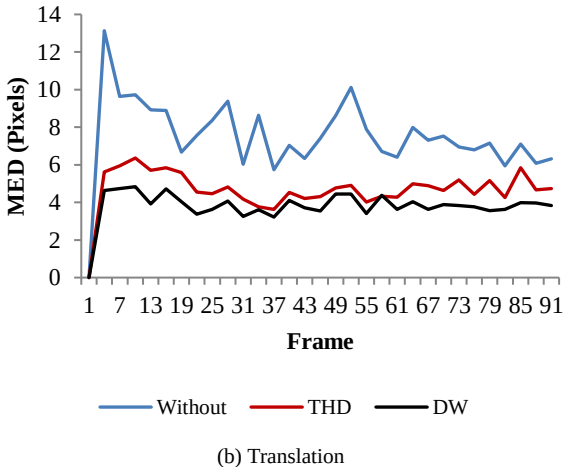
(b)

Figure 6: Final matched keypoint between previous and current image frames. (a) The corresponding points for all matched keypoints and (b) an example of nine corresponding keypoints for better visualization.

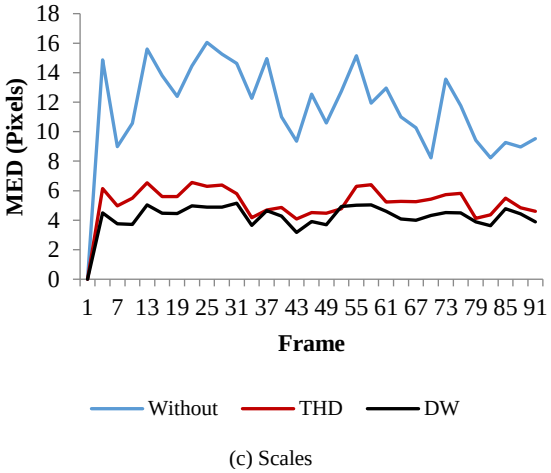
Besides, the graphs in Figure 7(a)-(e) show that the proposed distance criterion has successfully improved the MED performance. It was MED values decreased almost two times compared to the performance of matching algorithm without distance criterion. As shown in Figure 7(a)-(e), the DW criterion produced the lowest MED for each frame at all data set, inferring that the algorithm is capable of accurately matching the correct corresponding keypoints. Moreover, the DW MED performance showed a more consistent performance for the consecutive frames especially for rotation and translation movements and due scaling. This indicates that proposed distance criterion is robust for almost TDOD movement types.



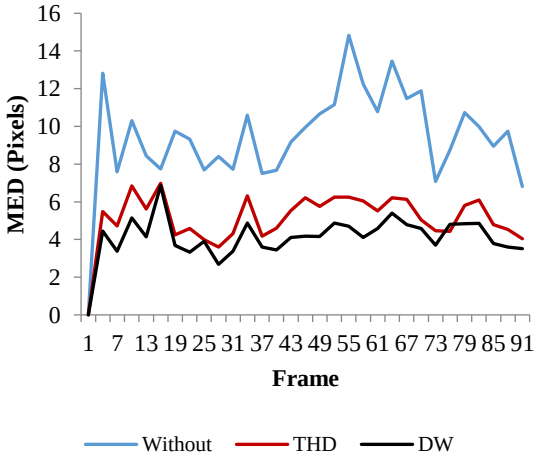
(a) Rotation



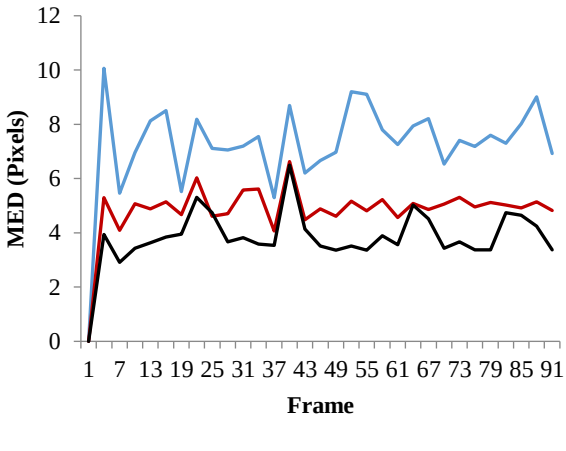
(b) Translation



(c) Scales



(d) Liver



(e) Heartbeat

Figure 7: Performance of distance criteria for each dataset

3) Comparison with Other Algorithms

In this section we are comparing the proposed matching algorithm with four benchmarks matching strategies: 1) threshold-based (THRES), 2) distance ratio (DR), 3) nearest neighbor threshold (NNT) and 4) nearest neighbor distance ratio (NNDR). The evaluation was done such that is not influenced by other factors such as the number of detected keypoints. The control parameters are as shown in Figure 8. The matching performances for the benchmark algorithms were measured by using Euclidean distance, with the threshold τ values are 0.4, 0.6 and 0.8. The average of MED for the threshold values is used as a comparison result.

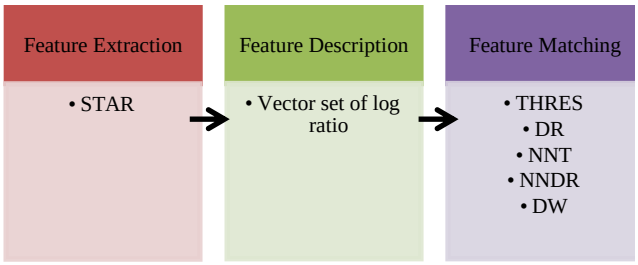


Figure 8: The diagram shows the process of validation for matching algorithms.

Figure 9 presents the MED performance of our DW algorithms compared to the four existing matching strategies. The THRES matching algorithm shows the worst MED performance compared to other matching algorithms. This is because the homogeneous texture leads to similar descriptor and yield the score of descriptor distance almost similar to each other. Hence, the THRES algorithm has very poor discriminative power. As a result, there are high chances of false matching.

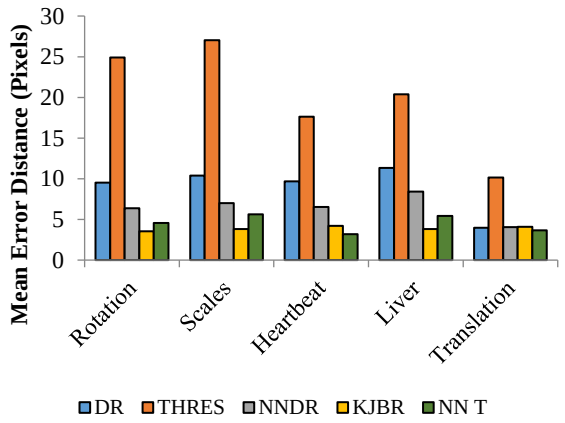


Figure 9: Comparison of matching performance of different types of algorithms

The DR algorithm shows a relatively better performance by drastically reducing the MED for each data set. The ratio method plays an important role in enhancing the differences between two similar descriptors. Meanwhile, the nearest neighbor techniques NNDR and NNT perform better than DR algorithm. These algorithms only process the most potential candidate keypoints and lead to the lower MED performance. The proposed DW algorithm gives the best performance by obtaining the lowest MED for each data set except for translation. However, the difference was very small and considered as insignificant. As a conclusion, the average MED performance (Figure 10) proved that our proposed matching algorithm performed better than existing matching.

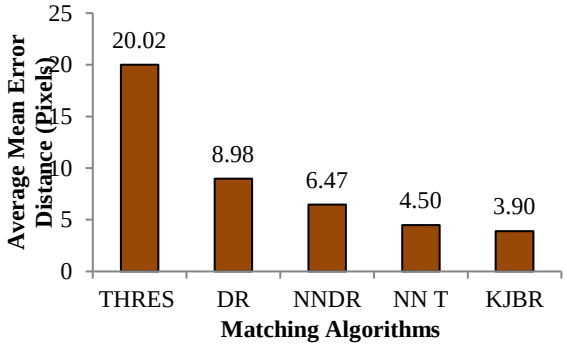


Figure 10: An average comparison of various matching algorithms.

IV. Conclusion

This paper discussed the improvement to the hypothesis test as a feature point matching algorithm for TDOD by incorporating multiple distance weighted criterions. This weightage plays important roles in providing the final correct decision. Test results show that the proposed algorithm gives the lowest MED compared to the threshold distance and without distance criterion. Besides, we have evaluated the performance of the proposed algorithm with four existing matching algorithms; THRES, DR, NNDR and NNT. Our proposed DW algorithm performed better compared to all these four algorithms. For future work, the matching performance can be improved by optimizing the cut-off distant used in this proposed criterion.

Acknowledgment

We would like to acknowledge funding from Universiti Malaysia Pahang (RDU1703233).

References

[1] A. Abdalbari, X. Huang, and J. Ren, "Endoscopy-MR Image Fusion for Image Guided Procedures," *Int. J. Biomed. Imaging*, 2013, 2013, pp. 23.

[2] A. M. Broehan, T. Rudolph, C. A. Amstutz, and J. H. Kowal, "Real-time multimodal retinal image registration for a computer-assisted laser photocoagulation system," *IEEE Trans. Biomed. Eng.* 58(10), 2011, pp. 2816–2824.

[3] A. Cifor, L. Risser, D. Chung, E. M. Anderson, and J. A. Schnabel, "Hybrid feature-based diffeomorphic registration for tumor tracking in 2-D liver ultrasound images," *IEEE Trans. Med. Imaging*. 32(9), 2013, pp. 1647–1656.

[4] A. Gupta, J. J. Little, and R. J. Woodham, "Using Line and Ellipse Features for Rectification of Broadcast Hockey Video," *Canadian Conference on Computer and Robot Vision*. 2011, pp. 32–39.

[5] A. Datta, Y. Sheikh, and T. Kanade, "Linearized motion estimation for articulated planes," *IEEE Trans. Pattern Anal. Mach. Intell*. 33(4), 2011, pp. 780–793.

[6] G. Salvietti, M. Malvezzi, G. Gioioso, and D. Prattichizzo, "On the use of homogeneous transformations to map human hand movements onto robotic hands," *IEEE International Conference on Robotics and Automation (ICRA)*. 2013, pp. 5352–5357.

[7] B. Zahneisen and T. Ernst, "Homogeneous coordinates in motion correction," *Magnetic resonance in medicine*. *75*(1), 2016, pp. 274-279.

[8] R. A. Karim, N. F. Zakaria, M. A. Zulkifley, M. M. Mustafa, I. Sagap, and N. H. Md Latar, "Telepointer technology in telemedicine: a review*," Biomed. Eng. Online*. 12(1), 2013, pp. 21.

[9] C.-H. Peng and C.-H. Cheng, "A panoramic endoscope design and implementation for Minimally Invasive Surgery," *IEEE International Symposium on Circuits and Systems (ISCAS)*. 2014, pp. 453–456.

[10] G. A. Puerto-Souza and G.-L. Mariottini, "A fast and accurate feature-matching algorithm for minimally-invasive endoscopic images," *IEEE Trans. Med. Imaging*. 32(7), 2013, pp. 1201–1214.

[11] Y. Wan, Q. Wu, and X. He, "Dense feature correspondence for video-based endoscope three-dimensional motion tracking," *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*. 2014, pp. 49–52.

[12] T. T. Nguyen, H. Jung, and D. Y. Lee. 2013. "Markerless tracking for augmented reality for image-guided Endoscopic Retrograde Cholangiopancreatography," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 7364–7.

[13] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points," *Proceedings Eighth IEEE International Conference on Computer Vision (ICCV 2001)*. 1, 2001, pp. 525–531.

[14] K. Mikolajczyk, "Scale & Affine Invariant Interest Point Detectors," *Int. J. Comput. Vis*. 60(1), 2004, pp. 63–86.

[15] C. Reimann and P. de Caritat, "New soil composition data for Europe and Australia: demonstrating comparability, identifying continental-scale processes and learning lessons for global geochemical mapping," *Sci. Total Environ*. 416, 2012, pp. 239–252.

[16] J. Palarea-Albaladejo and J. A. Martín-Fernánde, "Values below detection limit in compositional chemical data," *Analytica Chimica Acta*. 764(0), 2013, pp. 32–43.

[17] R. Tolosana-Delgado, "Uses and misuses of compositional data in sedimentology," *Sedimentary Geology*. 280, 2012, pp. 60–79.

[18] E. J. M. Carranza, "Analysis and mapping of geochemical anomalies using logratio-transformed stream sediment data with censored values," *Journal of Geochemical Exploration*. 110(2), 2011, pp. 167–185.

[19] C. Reimann, P. Filzmoser, K. Fabian, K. Hron, M. Birke, A. Demetriades, E. Dinelli, and A. Ladenberger, "The concept of compositional data analysis in practice--total major element concentrations in agricultural and grazing land soils of Europe," *Sci. Total Environ*. 426, 2012, pp. 196–210.

[20] R. Saaltink, J. Griffioen, G. Mol, and M. Birke, "Geogenic and agricultural controls on the geochemical composition of European agricultural soils," *J. Soils Sediments*. 14(1), 2014, pp. 121–137.

[21] X.-L. Sun, Y.-J. Wu, H.-L. Wang, Y.-G. Zhao, and G.-L. Zhang, "Mapping Soil Particle Size Fractions Using Compositional Kriging,

Cokriging and Additive Log-ratio Cokriging in Two Case Studies Math" *Geosci*. 46(4), 2013, pp. 429–443.

[22] M. L. C. Leite, "Applying compositional data methodology to nutritional epidemiology," *Stat. Methods Med. Res*., 25(6), 2016, pp. 3057-3065.

[23] R. A. Karim, M. M. Mustafa, and M. A. Zulkifley, "A log-ratio pair approach to endoscopic image matching," *IEEE Workshop on Statistical Signal Processing (SSP)*. 2014, pp. 185–188.

[24] M. Agrawal, K. Konolige, and M. R. Blas, "CenSurE : Center Surround Extremas for Realtime Feature Detection and Matching," *Computer Vision–ECCV 2008. Springer Berlin Heidelberg*. 2008, pp. 102-115.

[25] S. Ahn, B. Yi, Y. Suh, J. Kim, S. Lee, S. Shin, S. Shin, and E. Choi, "A feasibility study on the prediction of tumour location in the lung from skin motion," *Br. J. Radiol*. 77(919), 2004, pp. 588–596,

[26] K. M. Langen and D. T. L. Jones, "Organ motion and its management," *Int. J. Radiat. Oncol*. 50(1), 2001, pp. 265–278.

[27] M. A. Gates, R. A. Mekary, G. R. Chiu, E. L. Ding, G. A. Wittert, and A. B. Araujo, "Sex steroid hormone levels and body composition in men," *J. Clin. Endocrinol. Metab*. 98(6), 2013, pp. 2442–2450.

[28] M. D. Reisner, J. B. Grace, D. A. Pyke, and P. S. Doescher, "Conditions favouring Bromus tectorum dominance of endangered sagebrush steppe ecosystems," *J. Appl. Ecol*. 50(4), 2013, pp. 1039–1049.

[29] K. Das, M. Kudo, M. Kitano, H. Sakamoto, T. Komaki, T. Takagi, and K. Yamao, "Diagnostic value of endoscopic ultrasound-guided directional eFLOW in solid pancreatic lesions," *J. Med. Ultrason*. 40(3), 2013, pp. 211–218.

[30] I. Bodhinayake, M. Ottenhausen, M. A. Mooney, K. Kesavabhotla, P. Christos, J. T. Schwarz, and J. A. Boockvar, "Results and risk factors for recurrence following endoscopic endonasal transsphenoidal surgery for pituitary adenoma," *Clin. Neurol. Neurosurg*. 119, 2014, pp. 75–79.

[31] I. C. London, "Endoscopic Video Datasets." [Online]. Available: http://hamlyn.doc.ic.ac.uk/vision/. [Accessed: 21-Feb-2014]